动 物 学 研 究 2012，Feb. 33(1): 108−118 CN 53-1040/Q ISSN 0254-5853
Zoological Research DOI：10.3724/SP.J.1141.2012.01108

灵长类比较基因组学的研究进展

廖承红[1]，宿 兵[2,*]

(1. 海南大学 农学院，海口 570288；2. 中国科学院昆明动物研究所 遗传资源与进化国家重点实验室，昆明 650223)

摘要：随着人类和黑猩猩全基因组测序工作宣布完成，以及其他灵长类基因组测序工作的逐步开展，目前已经积累了大量的灵长类基因组数据，一个崭新的研究领域——灵长类比较基因组学应运而生。该文主要通过对人类和其他非人灵长类系统关系和基因组结构的比较，从系统进化、基因组结构和基因表达调控等方面评述该领域的研究进展，阐述人类、黑猩猩与其他非人灵长类之间的主要生物学差异，揭示人类进化的生物学机制。

关键词：灵长类；比较基因组；进化
中图分类号：Q349；Q959.848 **文献标志码**：A **文章编号**：0254-5853-(2012)01-0108-11

Research proceedings on primate comparative genomics

LIAO Cheng-Hong[1], SU Bing[2,*]

(1. *College of Agriculture*, *Hainan University*, *Haikou* 570288, *China*; 2. *State Key Laboratory of Genetic Resources and Evolution*, *Kunming Institute of Zoology*, *the Chinese Academy of Sciences*, *Kunming* 650223, *China*)

Abstract: With the accomplishment of genome sequencing of human, chimpanzee and other primates, there has been a great amount of primate genome information accumulated. Primate comparative genomics has become a new research field at current genome era. In this article, we reviewed recent progress in phylogeny, genome structure and gene expression of human and nonhuman primates, and we elaborated the major biological differences among human, chimpanzee and other non-human primate species, which is informative in revealing the mechanism of human evolution.

Key words: Primate; Comparative genomics; Evolution

2001年，人类基因组计划草图序列的完成标志着生命科学新纪元的开始，现在的人类基因组序列(build 35)近乎完整(International Human Genome Sequencing Consortium, 2004)，黑猩猩基因组草图已经初步完成(Chimpanzee Sequencing and Analysis Consortium, 2005)。进一步的工作正在推进，大猩猩、长臂猿等的“低覆盖度”版本基因组图也已完成。2007年，科学家破译了猕猴的基因组，这是继人类和黑猩猩之后，科学家破译出的第3种灵长类动物基因组。截至2009年，已有12种灵长类动物的基因组处于测序阶段(图1)(Marques-Bonet et al, 2009b)。2011年，科学家初步完成红毛猩猩(*Orangutan*)的基因组测序，红毛猩猩成为继人类和黑猩猩之后第3个基因组成功测序的猿类(Locke et al, 2011)。另外，婴猴和鼠狐猴的基因组草图也基本完成。

灵长类比较基因组学主要是基于人类和其他灵长类的基因组测序工作已经完成，通过比较基因组学的方法，探究灵长类与其他灵长类动物不同的基因组结构，以及人类与它们区别的特征，为研究人类基因的特有结构和功能提供重要线索，解开人类起源和进化的许多奥秘，也是揭示智慧秘密的重要一步。在基因组时代，从不断增加的测序基因组信息中，人们越来越熟练地解读其生物学意义，因而使人类的“亲戚”为人类确定自己的身份提供重要线索。同时，通过对灵长类基因组测序，了解人类疾病的遗传学基础，有助于人类疾病动物模型的建立。

本文将从系统进化、基因组结构和基因表达调控等方面，对灵长类比较基因学的研究进展进行综

收稿日期：2012-01-11；接受日期：2012-02-04
基金项目：国家自然科学基金重点项目(31130051)；海南省高等学校科学研究项目(Hjkj2011-11)；海南省自然科学基金(310036)
*通信作者(Corresponding author)，Tel/Fax:0871-5120212, E-mail: sub@mail.kiz.ac.cn

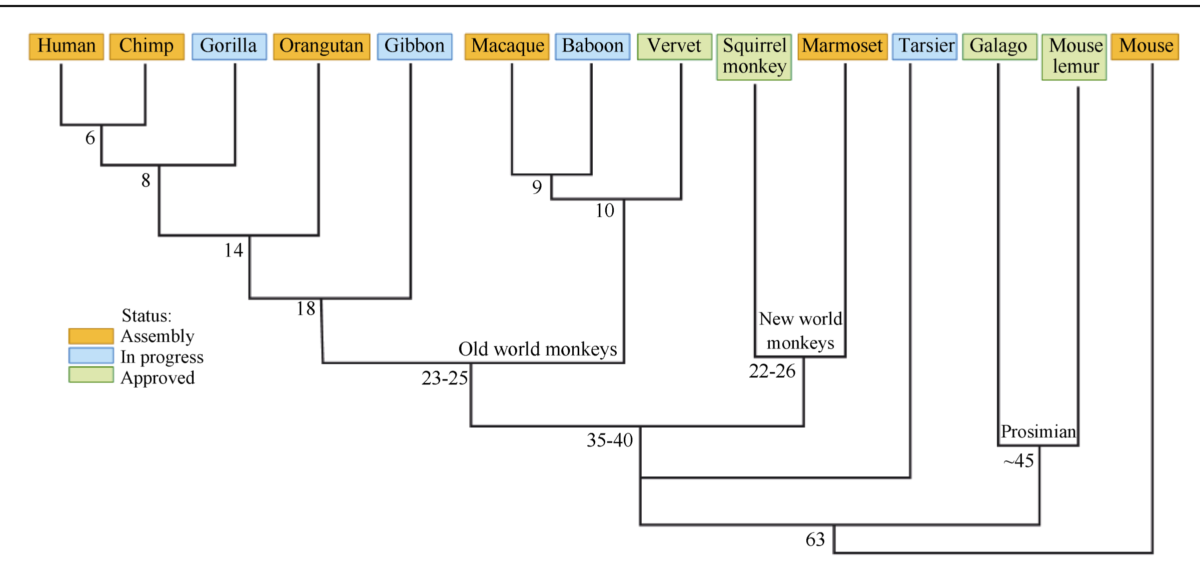


图 1　灵长类基因组测序的现状(Marques-Bonet et al, 2009b)

Fig. 1　Primate genome sequencing status(Marques-Bonet et al, 2009b)

分支上的数字表示各物种间的分歧时间(单位：百万年)。

The numbers on the branches are the divergence times of the species (million years).

述，初步阐述人类、黑猩猩与其它非人灵长类之间主要的遗传学差异，为揭示人类进化的机制提供线索。

1　灵长目的系统进化

1.1　系统进化关系

为了解基因组进化的结构和功能的特点，首先要掌握各个物种之间的关系，也就是它们的系统发育关系。目前，由于灵长类多个物种的基因组数据已测出，系统基因组学也随之出现。它是利用基因组测序结果来研究亲缘关系较近的物种之间的系统关系(Eisen & Fraser, 2003; Gomase & Tagore, 2009)。

目前灵长目系统进化关系研究已经比较清楚(图 2)，仅眼镜猴亚目的系统位置有一定的争议。因为研究发现在眼镜猴亚目、新大陆猴和旧大陆猴的相同位点上有 3 个相同的 Alu 序列，所以推测这 3 类亚目可能具有共同的祖先(Schmitz, et al, 2001)。总体上看，猿类是人类的近亲，其他灵长类是一些关系较远的“亲戚”，包括旧大陆猴、新大陆猴、眼镜猴和狐猴。

在猿类，非洲大猿与人类的关系比小猿和红毛猩猩更接近。但是，人类与黑猩猩和大猩猩之间的关系哪个更近呢？比较大猿和人类基因组数据时发现，有 18%的基因组区域在人类–黑猩猩分支上不一致，人类和大猿分支却相反(Chen & Li, 2001)。

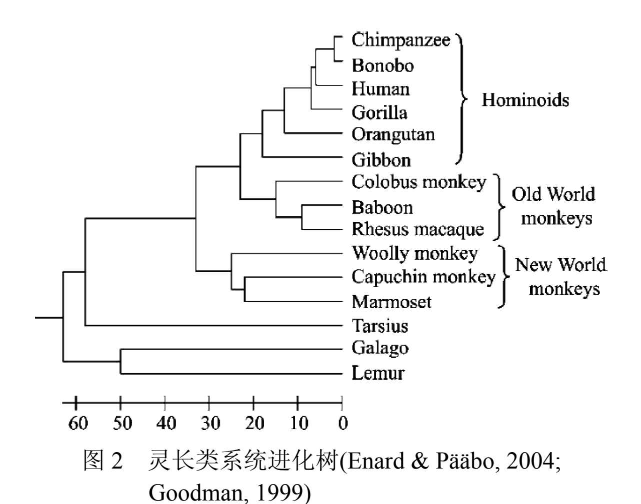


图 2　灵长类系统进化树(Enard & Pääbo, 2004; Goodman, 1999)

Fig. 2　Phylogenetic tree of primates (Enard & Pääbo, 2004; Goodman, 1999)

由于人类和黑猩猩、大猩猩之间亲缘关系较近，不能真实地反映三者之间的系统进化关系，只有利用 3 个物种以外的 1 个物种作为外群(outgroup)来进行比较才能解决这个问题。例如，与 3 个物种系统关系最近的是红毛猩猩，它的分支发生在更早以前，其基因组中没有任何一个区域受到平衡选择等特殊现象的影响，所以它作为外群可以真实反映人类和黑猩猩的系统进化关系。现在，人类、黑猩猩和红毛猩猩的基因组序列已经部分或全部完成，所以人们可以从基因组的角度来解释人类与黑猩猩之间的关系。通过对人类、黑猩猩和红毛猩猩 3 种基因组的对比，人们惊奇地发现红毛猩猩的基因组中

约 1%的序列与人类基因组更接近，而不是黑猩猩的基因组。这说明人类和黑猩猩的共同祖先经过几百万年分别进化，在这个过程中，黑猩猩因为某些原因失去了一部分共同祖先的 DNA，而人类则保留了这些 DNA(Hobolth et al, 2011)。

1.2 分歧时间

人们在分析灵长类准确分歧时间时，由于采用的手段(化石或 DNA 数据)、分析的数据来源(线粒体 DNA 或核 DNA)、利用的统计方法以及校正方法等的不同，都可能会给结果带来不确定性，因而，使得估计灵长类之间的分歧时间比确定系统进化关系更易产生分歧。例如，利用化石估计的狐猴与其他灵长类之间的分歧时间大约是 6 000 万年(Goodman et al, 1998)到 8 000 万年(Springer et al, 2003)之间，跨度时间较大。Glazko & Nei (2003)利用 2 种不同的化石作为校正标准以及多种统计方法检测，估计出人类和其他灵长类的分歧时间：人类和黑猩猩的分歧时间大约是 500～700 万年，大猩猩的是 600～800 万年，红毛猩猩是 1200～1500 万年，旧大陆猴是 2100～2500 万年，新大陆猴的是 3200～3600 万年。这些数据与过去的一些研究结果基本一致(Chen & Li, 2001; Goodman, 1999)。

在利用 DNA 数据估计关系较近的物种（如人类和黑猩猩）的分歧时间时，常容易忽略一个问题，即 DNA 序列的分歧时间不仅包括物种分开之后的时间，还包括 2 条序列从同一祖先群体中分化出来经历的时间(图 3)。如果物种分歧事件发生在较近时间，祖先群体较大，那么 DNA 序列的分歧时间就大于物种的分歧时间(Enard & Pääbo, 2004) 。

灭绝的灵长类物种的分子信息也可用于分析与人类的系统发育关系和分歧时间。尼安德特人就是一个典型的代表。尼安德特人(*Homo neanderthalensis*)生活于距今 50 万年前至距今约 3 万年前，分布于今天的欧洲、中东，以及亚洲的部分地区，而且现代人的祖先智人(*Homo sapiens*)与尼安德特人在历史上曾经都广泛分布在欧洲且共存过，因此，这两大人类集团很可能曾经相遇，并有机会杂交繁殖，这样，尼安德特人与现代人的关系最近。Krings et al (1997，2000)对尼安德特人和早期现代人的线粒体 DNA 进行了测序，估计现代人和尼安德特人的分歧时间为 50 万年左右，这与用考古学和古生物学估计的时间 35 万年接近。其中有 4 个尼安德特人的 mtDNA 序列被确定，表明

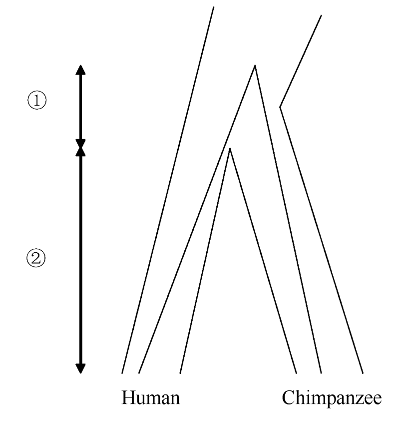


图 3 两条直系同源 DNA 序列的分歧时间
Fig. 3 The divergence time of two orthologous DNA sequences
1) 两条序列从同一祖先群体中分化出来经历的时间; 2) 物种分开之后的时间 (Enard & Pääbo, 2004)。
1) The time of the two sequences diverged from the common ancestral population; 2) the time after the two species separate (Enard & Pääbo, 2004).

某些存在于尼安德特人的 mtDNA 片段并不存在于现代人的 mtDNA 中，尼安德特人对早期现代人的基因没有显著贡献。但 Green et al (2010)将尼安德特人基因组测序结果与现代人基因组进行比较后发现，现代人与尼安德特人非常可能在小范围内发生过基因交流，时间可能是现代人走出非洲之后。

2 灵长类基因组的进化及机制

随着越来越多的灵长类 DNA 序列不断地积累增多，特别是黑猩猩、红毛猩猩和恒河猴基因组的测序工作基本完成，使灵长类比较基因组的研究有了丰富的数据资源。以下将从染色体进化、片段重复、插入缺失和单核苷酸多态等基因组的结构变化，蛋白质适应性进化和基因表达调控等方面对灵长类基因组进行比较，了解人类和非人灵长类之间从结构到功能的差异，在演化进程中(如从猿到人)所产生的变化及其遵循的规律。

2.1 染色体进化

灵长类进化过程中，染色体的变化相对保守，其中人类和红毛猩猩保持着与祖先更为相近的核型。相对而言，黑猩猩和大猩猩产生了一些物种特异的染色体重排(Müller & Wienberg, 2001)。人类和黑猩猩因常染色体重排导致 10 处不同核型：祖先 2 条染色体的端粒融合形成人的 2 号染色体和 9 处臂间倒位(Yunis & Prakash, 1982)，这些差异都没有导致功能上的明显变化(Enard & Pääbo, 2004)。研究发现 2 个姊妹黑猩猩物种——普通黑猩猩和倭黑猩猩

的近着丝点的倒位情况相同(Locke et al, 2003; Szamalek et al, 2006)。因此, 可以推断这些倒位发生在 2 种黑猩猩分开之前, 也就是 86～200 万年前(Yoder & Yang, 2000; Won & Hey, 2005), 而且重排染色体区域蛋白质进化的速度显著快于共线性染色体(colinear chromosomes)上的蛋白质(Vallender & Lahn, 2004a)。“重组抑制模型”认为在早期人类和黑猩猩还有可能杂交繁殖的时候, 染色体区域重排可能会阻碍人类和黑猩猩之间基因交流(Navarro & Barton, 2003a)。研究发现人类和黑猩猩在染色体重排区受到较强的正选择作用(Navarro & Barton, 2003b)，但是染色体重排对于灵长类的物种形成是否有重要作用, 还有待进一步的研究。

2.2 片段重复

根据所测基因组数据发现, 在灵长类基因组中存在大量的长片段的结构变化, 主要是由于片段重复等原因造成的。在人类基因组中存在很多重复片段(Li et al, 2001), 比较现有的全基因组数据, 人类基因组中片段重复数最高, 其覆盖率(约 5.3%)高于大鼠的基因组(约为 3%)和小鼠的基因组(约为 4.9%)(Bailey et al, 2004; Cheung et al, 2003; She et al, 2008)。这些重复序列位于相对较大的区域内(1～200 kb)的所谓的片段复制区, 至少每个单倍体基因组有两个拷贝, 且拷贝之间相似度高达 90%～100% (Samonte & Eichler, 2002)。

比较人类和大猿基因组, 发现人类和非洲大猿基因组中存在许多散在的重复序列, 说明这些重复序列已经发生在其共同祖先的基因组中(Marques-Bonet et al, 2009a), 并且发生在染色体内的片段重复数大于发生在染色体间的重复数(Zhang et al, 2005), 染色体内的大量片段重复发生在较近的 1000 万年前类人猿祖先的基因组中。相反, 染色体间的片段重复发生的时间更早, 大约在 2500 万年前, 即旧大陆猴和类人猿分开的时候(She et al, 2006)。在人类与黑猩猩分歧之后, 大约每百万年有 4～5 Mb 的片段重复发生, 33%的重复片段是人类特有的(Cheng et al, 2005)。重复片段往往群集于每条染色体的中部(着丝粒)或亚末端(端粒)附近(Bailey et al, 2002; Zhang et al, 2005)。与整个基因组相比, 重复片段明显趋向于聚集在这些区域(She et al, 2004), 并且由于人的亚端粒区基因密度相对较高(Flint et al, 1997), 因此，该区域广泛存在多态性(Daniels et al, 2001; Trask et al, 1998)。重复片段在人类的不同染色体之间分布差异很大, 其比例在 1%～14%之间变化。Y 染色体就是一个最极端的例子, 其重复片段，占总长度的 25%以上。

人类基因组的复杂性可能是由于进化过程中片段重复的结果(Stankiewicz et al, 2004)。由于重复片段的相似性都较高, 片段长度较长, 并且重复片段是来自于同一个物种的祖先序列, 而不是来自不同物种, 所以在基因组中较难对它们进行正确定位(Bailey et al, 2001)。而且, 如果把这些重复片段从基因组中忽略, 那么人类和黑猩猩之间的差异就会更低(Liu et al, 2003)。重复片段往往也是新基因或新基因家族形成的重要来源, 因此，片段重复可能在物种形成过程中对产生新的表型起到重要作用。在对 3 种灵长类(猕猴、红毛猩猩、黑猩猩)和人类基因组比较中发现片段重复引起的突变明显区别于其他遗传学上的突变(Marques-Bonet et al, 2009a), 并且这些重复在人类进化过程中起到重要的作用。人类基因组的高重复片段比率及与黑猩猩的较大差异, 表明人类的遗传物质经历了快速的功能变革和结构改变, 最终导致人类具有一些独特的特征。

2.3 插入和缺失

多年来，在比较研究人类和黑猩猩等非人灵长类基因组的过程中, 人们发现这些基因组之间的差异不仅仅局限于染色体重排的变化上, 还包括了许多微细结构上的变化(Britten 2002; Britten et al, 2003; Frazer et al, 2003; Kuroki et al, 2006; Liu et al, 2003)。这些变化包括了基因组内的插入和缺失, 存在约 3%，即 9 000 万个碱基插入或缺失的差异, 它们在人类和黑猩猩分开以后很大程度上改变了两者基因组的大小。不同的机制都可能导致插入或缺失的出现，如 DNA 复制、重组等。在人和黑猩猩 DNA 序列中，由于 DNA 复制引起的插缺在每 1000 个碱基内就能找到一个(Ebersberger et al, 2002)。

插缺在很大程度上影响人类和黑猩猩基因组大小的变化。基于人类 21 号染色体测序结果, Frazer et al (2003)用寡核苷酸芯片技术对人和黑猩猩等灵长类的 21 号染色体进行比较, 结果发现人类和黑猩猩的基因共有 57 处不同排列。从这个结果可以推断出, 在整个基因组的近 30 亿个基因当中, 两者存在大约 57 000 处不同排列, 并且在黑猩猩、红毛猩猩、恒河猴和新大陆猴有约 9%的缺失。比较人的 21 号染色体和黑猩猩对应的 22 号染色体, 330 万对碱基对中就有约 68 000 个插缺 (Watanabe et al,

2004)。这些插缺导致人第 21 号染色体要比黑猩猩第 22 号染色体长 400 kb, 其中 5%插缺可能引起蛋白序列改变或产生提前终止编码(Wetterbom et al, 2006)。这些研究说明了插缺在灵长类进化过程中的重要作用, 为解释人类和黑猩猩表型差异提供了信息。

2.4 单核苷酸的替换

人类和黑猩猩基因组进行排序比较, 结果显示黑猩猩和人类基因组差异性仅相差 1.23%, 比过去人们估计的还要低 (Chimpanzee Sequencing and Analysis Consortium, 2005)。整体上来看, 人类和黑猩猩是 2 个极其相近的物种, 两者之间的差异只相当于任意 2 个不同人之间基因组差异的 10 倍。它们核苷酸的变化程度受物种内多态性的很强影响, 如果不考虑物种内核苷酸多态性的话, 那么人类和黑猩猩的 DNA 序列差异性更低(~1%) (Chimpanzee Sequencing and Analysis Consortium, 2005)。

比较黑猩猩和人类 X 染色体上 10 kb 区域, 黑猩猩种内 DNA 序列的差异是人类的 3 倍(Kaessmann et al, 2001)。如果把人类、倭黑猩猩和黑猩猩(非洲东部、西部和中部)更大区域的 DNA 进行比较, 发现黑猩猩序列的种内差异是人类的 1.5 倍之多(Yu et al, 2003)。在中性位点上(即该位点发生改变并不影响生物体的适合度), 可以利用观察到的物种间替换率来估计突变率。大猿的每年突变率在进化过程中比较稳定, 旧大陆猴的突变率比大猿的高 30%, 狐猴的突变率是大猿的 2 倍(Liu et al, 2003), 小鼠的突变率是大猿的 5 倍(Waterston et al, 2002)。"世代假说"(generation time hypothesis)似乎能部分解释其差异性, 虽然现在大家对该假说还持争议。根据世代假说(Li et al, 1996), 大多数的生殖细胞产生的突变来自 DNA 的错误复制。世代时间越长, 每次参与复制的生殖细胞数量就越少, 那么积累固定下来的替换突变就越少。因此，世代时间长的物种, 分子进化速率(即进化钟运转速度)就比世代时间短的物种慢。在所有灵长类动物中, 人类的"进化钟"运转最慢(Elango et al, 2006), 也就是说人类基因组内单核苷酸变异更少, 基因组更稳定; 黑猩猩"进化钟"的运转仅比人类稍快一些, 但明显比大猩猩和猩猩慢, 从而进一步证明人类和黑猩猩的关系最近(Elango et al, 2006)。

人类和黑猩猩的单核苷酸替换率在基因组不同的范围又各不相同。比较人类 21 号染色体和黑猩猩对应染色体 22 号, 结果显示两者之间的差异(1.44%)大于基因组平均差异度(1.23%)(Watanabe et al, 2004)。其他染色体中, Y 染色体的核苷酸变化(1.78%)最大, 明显高于基因组整体差异性, 而 X 染色体的变化则最小(0.94%)(Hughes et al, 2005; Kuroki et al, 2006)。这可能是由于雄性生殖细胞突变率较快造成的 (Li et al, 2002; Makova & Li, 2002; Taylor et al, 2006)。另外, 人类和黑猩猩的常染色体的替换率也明显不同(Ebersberger et al, 2002; Watanabe et al, 2004)，其原因还不太明确。

另外，在考虑整个基因组核苷酸替换率时, CpG (CG 两联核苷酸常称作 CpG, 以表明连接两个核苷酸的磷酸二脂键)占 DNA 序列中的比例也是个重要因素 (Enard & Pääbo, 2004)。因为 CpG 二联核苷酸与一种重要的化学修饰——甲基化密切相关, 甲基化后的胞嘧啶特别容易发生突变(特别是突变成 TpG 和 CpA)。因此，在估计突变率时, 需要考虑 CpG 位点的变化 (Hellmann et al, 2003)。

2.5 蛋白质编码基因的适应性进化

人类与黑猩猩的蛋白质, 平均只差 2 个氨基酸, 约有 1/3 的蛋白质完全相同(Chimpanzee Sequencing and Analysis Consortium, 2005)。对于编码蛋白的基因, 非同义替换率 (nonsynonymous substitution rate, *Ka*)与同义替换率 (synonymous substitution rate, *Ks*)的比率(*Ka/Ks*)被广泛地用来检测基因在进化过程中受到怎样的选择作用(Miyata & Yasunaga, 1980；Li, 1993)。在这里, 同义替换被假定为是不改变生物的适合度的, 属于中性替换。如果 $Ka/Ks<1$, 说明该基因在进化过程中受到纯化选择(purifying selection)的作用, 即蛋白质功能受到限制; 若 $Ka/Ks = 1$, 则该基因被认为经历了一个中性的进化历程; 而如果 $Ka/Ks > 1$, 即某些氨基酸替换具有选择上的优势, 其固定的概率和速度要大于同义替换, 说明该基因在进化过程中受到了正选择(positive selection)作用。人和黑猩猩之间约 70%的氨基酸替代都是轻微有害的 (Hellmann et al, 2003)。然而, 也存在一部分替代是有利的。但是, 仅用 *Ka/Ks* 的值来决定整个编码区是否受到正选择, 往往灵敏度不够 (Ellegren, 2005; Kitano et al, 2004; Nielsen et al, 2005), 如有些 $Ka/Ks<1$ 的基因实际上在进化中受到强烈的正选择(Dorus et al, 2006)：蛋白质重要功能域的某个氨基酸改变, 也许不会很快表现出来(Andrés et al, 2004), 但蛋白质其他区域受到负选择

作用, 那么整个基因的 *Ka/Ks* 仍小于 1。考虑以上可能出现的情况, 现在已发展出一些新的方法研究某些分支或一些位点(branch-site)的正选择作用(Zhang et al, 2005)。CSAC 对人类和黑猩猩 13454 个同源基因对进行检验, 585 个基因在人类进化过程中可能受到正选择, *Ka/Ks*>1。但结果明显比其他 2 个类似的检验(Clark et al, 2003; Nielsen et al, 2005)得到的基因数少。Arbiza et al (2006)应用两分支－位点最大似然法(two branch-site maximum likehood)对人类、黑猩猩、大鼠、小鼠和狗的 13198 个同源基因进行检验, 结果显示人类和黑猩猩分别有 5%和 10%的基因受到正选择作用, 类似的结论也得到了更多的证据证明 (Bakewell et al, 2007)。人类受正选择的基因之所以比黑猩猩的少, 可能是进化过程中人类有效群体较小的缘故 (Chen & Li, 2001)。

研究发现受到正选择的基因有一定的功能偏好性, 人类基因组中受到正选择的基因主要是与免疫、生殖、神经功能以及一些与特殊环境相关的基因(Miller et al, 2004; Vallender & Lahn, 2004b)。这些基因中, 特别是免疫和生殖相关的基因不仅在人类，而且在黑猩猩都受到选择作用 (Arbiza et al, 2006), 而与脑功能/神经发育相关基因的适应性进化从一开始就与人类出现特有的特征相联系。有些基因在灵长类进化过程中受到正选择作用, 并且在现代人群中也仍然还存在选择信号, 如影响脑容量大小的两个基因 *MCPH1* (*microcephalin*)和 *ASPM* (abnormal spindle-like microcephaly-associated), 这 2 个基因的编码区发生某些突变, 都会导致原发性小脑症(primary microcephaly, MCPH)——脑容量是正常人的 1/3, 但脑部沟回等结构却基本正常(Bond et al, 2002; Jackson et al, 2002; Roberts et al, 2002)。*ASPM* 在灵长类的进化和人类起源过程中受到显著的正选择作用, *MCPH1* 的进化模式也是类似的(Evans et al, 2004a, b; Evans et al, 2005; Kouprina et al, 2004; Wang & Su, 2004; Zhang, 2003), 这 2 个基因可能在人类起源过程中对脑容量的增大发挥一定的作用。最近研究发现这两个基因在现代人群中仍然受到正选择作用, 说明现代人的脑还处于继续进化的过程中 (Evans et al, 2005; Mekel-Bobrov et al, 2005)。在人类起源过程中受到达尔文正选择的基因, 还可能是由于编码区的少数几个氨基酸的改变, 导致了蛋白质功能的改变, 进而引起表型的巨大变异。一个典型的例子就是对与人类语言功能相关的基因 *FOXP2* 的序列比较研究(Enard et al, 2002b; Lai et al, 2001)。

Dorus et al (2004)通过比较灵长类(人类和猕猴)和啮齿类(大鼠和小鼠)的 214 个神经系统相关基因的错义突变和同义突变比率(*Ka/Ks*)发现, 神经系统相关基因特别是脑发育相关基因在人类起源过程中有加速进化的趋势, 表明这一类基因可能受到正选择。进一步比较人和猕猴 (以松鼠猴作外群)、人和黑猩猩 (以猕猴作外群)的序列后发现, 神经系统相关基因在人这一支系 *Ka/Ks* 较大; 但是研究发现, 当人类和黑猩猩进行比较, 脑中表达量极高的那些基因在人类这一支很少或不受到正选择作用(Nielsen et al, 2005); 脑部表达的 2633 个基因, 也仅有 47 个基因在人类进化中受到正选择, 而黑猩猩没有 (Yu et al, 2006)。这些结果显示, 在进化过程中那些在脑的某些部位特异表达或是表达量较低的基因, 或是那些表达量发生改变的基因, 可能与人类脑发育和认知能力密切相关。

2.6　基因表达调控

由于基因组 DNA 序列上, 特别是蛋白质水平上的高度相似难以解释人类和黑猩猩之间认知能力上的巨大差异, 因此，人们试图从基因表达水平上来探讨这个问题(Carroll, 2005; King & Wilson, 1975)。有一种观点认为人和其他动物认知能力上的巨大差异是由于基因表达水平的差异引起的(Enard et al, 2002a; King & Wilson, 1975; Marvanova et al, 2003; Normile, 2001)。

近年来, 研究者采用基因芯片技术(microarray)研究基因的表达差异。利用该技术比较人类和非人灵长类的基因表达谱, 揭示在人类和非人灵长类之间差异表达的基因和基因家族, 以及相关代谢途径中的基因群, 已经成为了近年来研究的一个热点(Cáceres et al, 2003; Enard et al, 2002a; Gilad et al, 2006; Karaman et al, 2003; Khaitovich et al, 2005; Marvanova et al, 2003)。这些研究结果发现, 人类和黑猩猩的基因表达谱存在物种特异性和组织特异性的特点, 而且这些差异并不是因为环境因素的差异随机产生的, 而是有一定的遗传学基础的。利用芯片技术还检测到在衰老过程中人类和黑猩猩脑组织许多基因表达发生变化, 但是两者变化的模式明显不同。这提示在不同物种的整个基因组中, 可能存在着一些包括与衰老相关的基因表达量快速进化的改变 (Fraser et al, 2005)。比较人类和黑猩猩

各组织基因表达情况，发现两者存在表达差异基因数最多的组织是睾丸 (Khaitovich et al, 2005)。相反，存在差异基因数最少的是脑组织，这可能与在高级灵长类中脑功能相关基因的表达受到功能限制有关 (Enard et al, 2002a; Khaitovich et al, 2005)；但是，如果在人类和黑猩猩脑组织中表达存在差异的基因，往往在人类脑组织中的表达量要高些(Cáceres et al, 2003; Enard et al, 2002a)，并且这些在人脑中上调的基因往往是人脑组织特异表达的基因(Gilad et al, 2006)。然而，黑猩猩脑中有些与智力相关如快速识别眼−手信号等的基因表达增强(Inoue & Matsuzawa, 2007)。通过比较还发现了一些规律，比如性染色体、重复片段和染色体重排的区域在两个物种之间表达差异较大(Marquès-Bonet et al, 2004)。

造成基因表达存在差异的原因可能是基因的调控区序列差异、调控起始受到不同的控制、RNA 加工和翻译、染色体结构的不同修饰、选择性剪切以及 microRNA 的调控等。基因启动子序列的变化在某种程度上影响人类个体的基因转录水平(Bray et al, 2003; Hoogendoorn et al, 2003; Trinklein et al, 2003)，可能是启动子序列变化导致顺式(cis-)和反式(trans-)调控因子的结合能力改变(Pilpel et al, 2001)。个体间基因表达差异受到多个 cis-和 trans-调控因子的影响，物种间基因表达差异也可能受到多种遗传因素的作用。人们在研究高等灵长类基因表达差异时，发现在人类这一支系中转录因子比人类其它蛋白的进化速度要快。转录因子的加速进化暗示，为什么发生在关键位置上的如此少的遗传学改变，能通过改变转录因子结合能力而多效地影响许多不同基因的表达模式。因此，我们可以推测引起人类和黑猩猩之间基因表达差异的最主要因素是人类转录因子的快速进化。为了更充分地了解灵长类基因转录体的进化，还需要对更多的个体和组织进行研究。现在使用的芯片杂交技术往往因为灵长类之间序列差异受到影响，进而影响了结果的运算处理，如果能对各个灵长类设计其特异的芯片的话，将大大改变现状。

探究基因表达差异的原因，人们在比较人类和黑猩猩 DNA 甲基化情况时，还发现人类脑组织中甲基化程度远大于黑猩猩脑组织，其他人体组织，如肝脏和淋巴细胞也出现高水平的甲基化，但并不如脑组织中的变化大 (Enard et al, 2004)。目前还没有证据证明甲基化可能引起以往人们研究中观察到的基因表达差异 (Enard & Pääbo, 2004)。

最近，许多研究发现物种间基因表达差异还受到 microRNA 的调控。microRNA 可以通过与 mRNA 的 3´UTR 的结合降解该分子，是一种对基因表达具有重要调控作用的非编码 RNA，对功能基因 3´UTR 的进化有明显的影响 (Stark et al, 2005)。在人类和其他灵长类大脑 miRNAs 对比中，还发现许多灵长类进化产生的新 miRNAs 不是保守的，这说明这些 miRNA 起源比较晚，受到强烈的正选择，存在快速的进化 (Zhang et al, 2007)，而且一些 miRNAs 是物种特异性的，但其通过复制在物种中出现了扩增，这些数据说明 miRNAs 的进化还在进行，并且除了一些古老的高保守性的 miRNAs，也出现了一些新 miRNAs。

通过对人类和非人灵长类基因组的比较，除了发现人类特异的突变、人类特有(无)的基因、受正选择的基因以及表达差异的基因外，还观察到一些人类特异的重复和假基因化事件(Fortna et al, 2004; Wang et al, 2006)，这些差异都有可能在人类起源过程中发挥作用，通过对其研究有助于进一步揭示人类起源过程的遗传本质。

3 小结和展望

目前，除了黑猩猩、猕猴和红毛猩猩全基因组测序基本完成外，其他灵长类基因组测序还在进行中。随着这些基因组测序工作的完成，将给灵长类比较基因组学带来更丰富的数据资源，对人们了解灵长类的进化，以及更好地认识人类本身带来广阔的前景。在此基础上，建立合适的灵长类动物模型，研究基因组中插缺、基因失活和基因转座等现象；通过有效整合各个基因组基因表达、DNA 序列变化和功能变化的数据，有助于充分认识有关灵长类特别是人类表型改变的遗传学机制等一系列问题。各灵长类物种基因组测序的完成将是比较基因组学的重大机遇，不仅对基因组学，而且对神经科学、行为科学、分子生物学、遗传学、发育生物学等都有重大的影响。在此基础上，人类与非人灵长类基因组比较工作将全面推进，解释人类进化的许多疑问，为揭示人类起源之谜翻开新的篇章。

参考文献：

Andrés AM, Soldevila M, Navarro A, Kidd KK, Oliva B, Bertranpetit J. 2004. Positive selection in MAOA gene is human exclusive: determination of the putative amino acid change selected in the human lineage[J]. *Hum Genet*, **115**(5): 377-386.

Arbiza L, Dopazo J, Dopazo H. 2006. Positive selection, relaxation, and acceleration in the evolution of the human and chimp genome[J]. *PLoS Comput Biol*, **2**(4): e38.

Bailey JA, Church DM, Ventura M, Rocchi M, Eichler EE. 2004. Analysis of segmental duplications and genome assembly in the mouse[J]. *Genome Res*, **14**(5): 789-801.

Bailey JA, Gu ZP, Clark RA, Reinert K, Samonte RV, Schwartz S, Adams MD, Myers EW, Li PW, Eichler EE. 2002. Recent segmental duplications in the human genome[J]. *Science*, **297**(5583): 1003-1007.

Bailey JA, Yavor AM, Massa HF, Trask BJ, Eichler EE. 2001. Segmental duplications: organization and impact within the current human genome project assembly[J]. *Genome Res*, **11**(6): 1005-1017.

Bakewell MA, Shi P, Zhang JZ. 2007. More genes underwent positive selection in chimpanzee evolution than in human evolution[J]. *Proc Natl Acad Sci USA*, **104**(18): 7489-7494.

Bond J, Roberts E, Mochida GH, Hampshire DJ, Scott S, Askham JM, Springell K, Mahadevan M, Crow YJ, Markham AF, Walsh CA, Woods CG. 2002. *ASPM* is a major determinant of cerebral cortical size[J]. *Nat Genet*, **32**(2): 316-320.

Bray NJ, Buckland PR, Owen MJ, O'Donovan MC. 2003. Cis-acting variation in the expression of a high proportion of genes in human brain[J]. *Hum Genet*, **113**(2): 149-153.

Britten RJ. 2002. Divergence between samples of chimpanzee and human DNA sequences is 5%, counting indels[J]. *Proc Natl Acad Sci USA*, **99**(21): 13633-13635.

Britten RJ, Rowen L, Williams J, Cameron RA. 2003. Majority of divergence between closely related DNA samples is due to indels[J]. *Proc Natl Acad Sci USA*, **100**(8): 4661-4665.

Cáceres M, Lachuer J, Zapala MA, Redmond JC, Kudo L, Geschwind DH, Lockhart DJ, Preuss TM, Barlow C. 2003. Elevated gene expression levels distinguish human from non-human primate brains[J]. *Proc Natl Acad Sci USA*, **100**(22): 13030-13035.

Carroll SB. 2005. Evolution at two levels: on genes and form[J]. *PLoS Biol*, **3**(7): e245.

Chen FC, Li WH. 2001. Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees[J]. *Am J Hum Genet*, **68**(2): 444-456.

Cheng Z, Ventura M, She XW, Khaitovich P, Graves T, Osoegawa K, Church D, DeJong P, Wilson RK, Pääbo S, Rocchi M, Eichler EE. 2005. A genome-wide comparison of recent chimpanzee and human segmental duplications[J]. *Nature*, **437**(7055): 88-93.

Cheung J, Wilson MD, Zhang JJ, Khaja R, MacDonald JR, Heng HH, Koop BF, Scherer SW. 2003. Recent segmental and gene duplications in the mouse genome[J]. *Genome Biol*, **4**(8): R47.

Chimpanzee Sequencing and Analysis Consortium. 2005. Initial sequence of the chimpanzee genome and comparison with the human genome[J]. *Nature*, **437** (7055): 69-87

Clark AG, Glanowski S, Nielsen R, Thomas PD, Kejariwal A, Todd MA, Tanenbaum DM, Civello D, Lu F, Murphy B, Ferriera S, Wang G, Zheng XG, White TG, Sninsky JJ, Adams MD, Cargill M. 2003. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios[J]. *Science*, **302**(5652): 1960-1963.

Daniels RJ, Peden JF, Lloyd C, Horsley SW, Clark K, Tufarelli C, Kearney L, Buckle VJ, Doggett NA, Flint J, Higgs DR. 2001. Sequence, structure and pathology of the fully annotated terminal 2 Mb of the short arm of human chromosome 16[J]. *Hum Mol Genet*, **10**(4): 339-352.

Dorus S, Anderson JR, Vallender EJ, Gilbert SL, Zhang L, Chemnick LG, Ryder OA, Li W, Lahn BT. 2006. *Sonic Hedgehog*, a key development gene, experienced intensified molecular evolution in primates[J]. *Hum Mol Genet*, **15**(13): 2031-2037.

Dorus S, Vallender EJ, Evans PD, Anderson JR, Gilbert SL, Mahowald M, Wyckoff GJ, Malcom CM, Lahn BT. 2004. Accelerated evolution of nervous system genes in the origin of *Homo sapiens*[J]. *Cell*, **119**(7): 1027-1040.

Ebersberger I, Metzler D, Schwarz C, Pääbo S. 2002. Genomewide comparison of DNA sequences between humans and chimpanzees[J]. *Am J Hum Genet*, **70**(6): 1490-1497.

Eisen JA, Fraser CM. 2003. Phylogenomics: intersection of evolution and genomics[J]. *Science*, **300**(5626): 1706-1707.

Elango N, Thomas JW, Yi SV. 2006. Variable molecular clocks in hominoids[J]. *Proc Natl Acad Sci USA*, **103**(5): 1370-1375.

Ellegren H. 2005. Evolution: natural selection in the evolution of humans and chimps[J]. *Curr Biol*, **15**(22): R919-R922.

Enard W, Fassbender A, Model F, Ador-jan P, Paabo S, Olek A. 2004.Differences in DNA methylation patterns between humans and chimpanzees[J].*Curr.Biol.***14**(4):R148–149

Enard W, Khaitovich P, Klose J, Zöllner S, Heissig F, Giavalisco P, Nieselt-Struwe K, Muchmore E, Varki A, Ravid R, Doxiadis GM, Bontrop RE, Pääbo S. 2002a. Intra- and interspecific variation in primate gene expression patterns[J]. *Science*, **296**(5566): 340-343.

Enard W, Przeworski M, Fisher SE, Lai CS, Wiebe V, Kitano T, Monaco AP, Pääbo S. 2002b. Molecular evolution of *FOXP2*, a gene involved in speech and language[J]. *Nature*, **418**(6900): 869-872.

Enard W, Pääbo S. 2004. Comparative primate genomics[J]. *Annu Rev Genomics Hum Genet*, **5**: 351-378.

Evans PD, Anderson JR, Vallender EJ, Choi SS, Lahn BT. 2004a. Reconstructing the evolutionary history of *microcephalin*, a gene controlling human brain size[J]. *Hum Mol Genet*, **13**(11): 1139-1145.

Evans PD, Anderson JR, Vallender EJ, Gilbert SL, Malcom CM, Dorus S, Lahn BT. 2004b. Adaptive evolution of *ASPM*, a major determinant of cerebral cortical size in humans[J]. *Hum Mol Genet*, **13**(5): 489-494.

Evans PD, Gilbert SL, Mekel-Bobrov N, Vallender EJ, Anderson JR, Vaez-Azizi LM, Tishkoff SA, Hudson RR, Lahn BT. 2005. *Microcephalin*, a gene regulating brain size, continues to evolve adaptively in humans[J]. *Science*, **309**(5741): 1717-1720.

Flint J, Thomas K, Micklem G, Raynham H, Clark K, Doggett NA, King A, Higgs DR. 1997. The relationship between chromosome structure and function at a human telomeric region[J]. *Nat Genet*, **15**(3): 252-257.

Fortna A, Kim Y, MacLaren E, Marshall K, Hahn G, Meltesen L, Brenton M, Hink R, Burgers S, Hernandez-Boussard T, Karimpour-Fard A, Glueck D, McGavran L, Berry R, Pollack J, Sikela JM. 2004. Lineage-specific gene duplication and loss in human and great ape evolution[J]. *PLoS Biol*, **2**(7): E207.

Fraser HB, Khaitovich P, Plotkin JB, Pääbo S, Eisen MB. 2005. Aging and gene expression in the primate brain[J]. *PLoS Biol*, **3**(9): e274.

Frazer KA, Chen XY, Hinds DA, Pant PV, Patil N, Cox DR. 2003. Genomic DNA insertions and deletions occur frequently between humans and nonhuman primates[J]. *Genome Res*, **13**(3): 341-346.

Gilad Y, Oshlack A, Smyth GK, Speed TP, White KP. 2006. Expression

profiling in primates reveals a rapid evolution of human transcription factors[J]. *Nature*, **440**(7081): 242-245.

Glazko GV, Nei M. 2003. Estimation of divergence times for major lineages of primate species[J]. *Mol Biol Evol*, **20**(3): 424-434.

Gomase VS, Tagore S. 2009. Phylogenomics: evolution and genomics intersection[J]. *Int J Bioinform Res Appl*, **5**(5): 548-563.

Goodman M. 1999. The genomic record of Humankind's evolutionary roots[J]. *Am J Hum Genet*, **64**(1): 31-39.

Goodman M, Porter CA, Czelusniak J, Page SL, Schneider H, Shoshani J, Gunnell G, Groves CP. 1998. Toward a phylogenetic classification of Primates based on DNA evidence complemented by fossil evidence[J]. *Mol Phylogenet Evol*, **9**(3): 585-598.

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai WW, Fritz MHY, Hansen NF, Durand EY, Malaspinas AS, Jensen JD, Marques-Bonet T, Alkan C, Prüfer K, Meyer M, Burbano HA, Good JM, Schultz R, Aximu-Petri A, Butthof A, Höber B, Höffner B, Siegemund M, Weihmann A, Nusbaum C, Lander ES, Russ C, Novod N, Affourtit J, Egholm M, Verna C, Rudan P, Brajkovic D, Kucan Ž, Gušic I, Doronichev VB, Golovanova LV, Lalueza-Fox C, de la Rasilla M, Fortea J, Rosas A, Schmitz RW, Johnson PLF, Eichler EE, Falush D, Birney E, Mullikin JC, Slatkin M, Nielsen R, Kelso J, Lachmann M, Reich D, Pääbo S. 2010. A draft sequence of the Neandertal genome[J]. *Science*, **328**(5979): 710-722.

Hellmann I, Zöllner S, Enard W, Ebersberger I, Nickel B, Paabo S. 2003. Selection on human genes as revealed by comparisons to chimpanzee cDNA[J]. *Genome Res*, **13**(5): 831-837.

Hobolth A, Dutheil JY, Hawks J, Schierup MH, Mailund T. 2011. Incomplete lineage sorting patterns among human, chimpanzee, and orangutan suggest recent orangutan speciation and widespread selection[J]. *Genome Res*, **21**(3): 349-356.

Hoogendoorn B, Coleman SL, Guy CA, Smith K, Bowen T, Buckland PR, O'Donovan MC. 2003. Functional analysis of human promoter polymorphisms[J]. *Hum Mol Genet*, **12**(18): 2249-2254.

Hughes JF, Skaletsky H, Pyntikova T, Minx PJ, Graves T, Rozen S, Wilson RK, Page DC. 2005. Conservation of Y-linked genes during human evolution revealed by comparative sequencing in chimpanzee[J]. *Nature*, **437**(7055): 100-103.

Inoue S, Matsuzawa T. 2007. Working memory of numerals in chimpanzees[J]. *Curr Biol*, **17**(23): R1004-R1005.

International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome[J]. *Nature.* **409** (6822): 860–921.

Jackson AP, Eastwood H, Bell SM, Adu J, Toomes C, Carr IM, Roberts E, Hampshire DJ, Crow YJ, Mighell AJ, Karbani G, Jafri H, Rashid Y, Mueller RF, Markham AF, Woods CG. 2002. Identification of microcephalin, a protein implicated in determining the size of the human brain[J]. *Am J Hum Genet*, **71**(1): 136-142.

Kaessmann H, Wiebe V, Weiss G, Pääbo S. 2001. Great ape DNA sequences reveal a reduced diversity and an expansion in humans[J]. *Nat Genet*, **27**(2): 155-156.

Karaman MW, Houck ML, Chemnick LG, Nagpal S, Chawannakul D, Sudano D, Pike BL, Ho VV, Ryder OA, Hacia JG. 2003. Comparative analysis of gene-expression patterns in human and African great ape cultured fibroblasts[J]. *Genome Res*, **13**(7): 1619-1630.

Khaitovich P, Hellmann I, Enard W, Nowick K, Leinweber M, Franz H, Weiss G, Lachmann M, Pääbo S. 2005. Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees[J]. *Science*, **309**(5742): 1850-1854.

King MC, Wilson AC. 1975. Evolution at two levels in humans and chimpanzees[J]. *Science*, **188**(4184): 107-116.

Kitano T, Liu YH, Ueda S, Saitou N. 2004. Human-specific amino acid changes found in 103 protein-coding genes[J]. *Mol Biol Evol*, **21**(5): 936-944.

Kouprina N, Pavlicek A, Mochida GH, Solomon G, Gersch W, Yoon YH, Collura R, Ruvolo M, Barrett JC, Woods CG, Walsh CA, Jurka J, Larionov V. 2004. Accelerated evolution of the ASPM gene controlling brain size begins prior to human brain expansion[J]. *PLoS Biol*, **2**(5): E126.

Krings M, Capelli C, Tschentscher F, Geisert H, Meyer S, von Haeseler A, Grossschmidt K, Possnert G, Paunovic M, Pääbo S. 2000. A view of Neandertal genetic diversity[J]. *Nat Genet*, **26**(2): 144-146.

Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Pääbo S. 1997. Neandertal DNA sequences and the origin of modern humans[J]. *Cell*, **90**(1): 19-30.

Kuroki Y, Toyoda A, Noguchi H, Taylor TD, Itoh T, Kim DS, Kim DW, Choi SH, Kim IC, Choi HH, Kim YS, Satta Y, Saitou N, Yamada T, Morishita S, Hattori M, Sakaki Y, Park HS, Fujiyama A. 2006. Comparative analysis of chimpanzee and human Y chromosomes unveils complex evolutionary pathway[J]. *Nat Genet*, **38**(2): 158-167.

Lai CSL, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. 2001. A forkhead-domain gene is mutated in a severe speech and language disorder[J]. *Nature*, **413**(6855): 519-523.

Li WH. 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitution[J]. *J Mol Evol*, **36**(1): 96-99.

Li WH, Ellsworth DL, Krushkal J, Chang BHJ, Hewett-Emmett D. 1996. Rates of nucleotide substitution in primates and rodents and the generation-time effect hypothesis[J]. *Mol Phylogenet Evol*, **5**(1): 182-187.

Li WH, Gu ZL, Wang HD, Nekrutenko A. 2001. Evolutionary analyses of the human genome[J]. *Nature*, **409**(6822): 847-849.

Li WH, Yi S, Makova K. 2002. Male-driven evolution[J]. *Curr Opin Genet Dev*, **12**(6): 650-656.

Liu G, Zhao SY, Bailey JA, Sahinalp SC, Alkan C, Tuzun E, Green ED, Eichler EE. 2003. Analysis of primate genomic variation reveals a repeat-driven expansion of the human genome[J]. *Genome Res*, **13**(3): 358-368.

Locke DP, Archidiacono N, Misceo D, Cardone MF, Deschamps S, Roe B, Rocchi M, Eichler EE. 2003. Refinement of a chimpanzee pericentric inversion breakpoint to a segmental duplication cluster[J]. *Genome Biol*, **4**(8): R50.

Locke DP, Hillier LW, Warren WC, Worley KC, Nazareth LV, Muzny DM, Yang SP, Wang Z, Chinwalla AT, Minx P, Mitreva M, Cook L, Delehaunty KD, Fronick C, Schmidt H, Fulton LA, Fulton RS, Nelson JO, Magrini V, Pohl C, Graves TA, Markovic C, Cree A, Dinh HH, Hume J, Kovar CL, Fowler GR, Lunter G, Meader S, Heger A, Ponting CP, Marques-Bonet T, Alkan C, Chen L, Cheng Z, Kidd JM, Eichler EE, White S, Searle S, Vilella AJ, Chen Y, Flicek P, Ma J, Raney B, Suh B, Burhans R, Herrero J, Haussler D, Faria R, Fernando O, Darré F, Farré D, Gazave E, Oliva M, Navarro A, Roberto R, Capozzi O, Archidiacono N, Della Valle G, Purgato S, Rocchi M, Konkel MK, Walker JA, Ullmer B, Batzer MA, Smit AF, Hubley R, Casola C, Schrider DR, Hahn MW, Quesada V, Puente XS, Ordoñez GR, López-Otín C, Vinar T, Brejova B, Ratan A, Harris RS, Miller W, Kosiol C, Lawson HA, Taliwal V, Martins AL, Siepel A, Roychoudhury A, Ma X, Degenhardt J, Bustamante CD, Gutenkunst RN, Mailund T, Dutheil JY, Hobolth A, Schierup MH, Ryder OA, Yoshinaga Y, de Jong PJ, Weinstock GM, Rogers J, Mardis ER, Gibbs RA, Wilson RK. 2011. Comparative and demographic analysis of orang-utan genomes[J]. *Nature*, **469**(7331): 529-533.

Makova KD, Li WH. 2002. Strong male-driven evolution of DNA sequences in humans and apes[J]. *Nature*, **416**(6881): 624-626.

Marquès-Bonet T, Cáceres M, Bertranpetit J, Preuss TM, Thomas JW,

Navarro A. 2004. Chromosomal rearrangements and the genomic distribution of gene-expression divergence in humans and chimpanzees[J]. *Trends Genet*, **20**(11): 524-529.

Marques-Bonet T, Kidd JM, Ventura M, Graves TA, Cheng Z, Hillier LW, Jiang Z, Baker C, Malfavon-Borja R, Fulton LA, Alkan C, Aksay G, Girirajan S, Siswara P, Chen L, Cardone MF, Navarro A, Mardis ER, Wilson RK, Eichler EE. 2009a. A burst of segmental duplications in the genome of the African great ape ancestor[J]. *Nature*, **457**(7231): 877-881.

Marques-Bonet T, Ryder OA, Eichler EE. 2009b. Sequencing primate genomes: what have we learned?[J]. *Annu Rev Genomics Hum Genet*, **10**(1): 355-386.

Marvanova M, Menager J, Bezard E, Bontrop RE, Pradier L, Wong G. 2003. Microarray analysis of nonhuman primates: validation of experimental models in neurological disorders[J]. *Faseb J*, **17**(8): 929-931.

Mekel-Bobrov N, Gilbert SL, Evans PD, Vallender EJ, Anderson JR, Hudson RR, Tishkoff SA, Lahn BT. 2005. Ongoing adaptive evolution of *ASPM*, a brain size determinant in *Homo sapiens*[J]. *Science*, **309**(5741): 1720-1722.

Miller W, Makova KD, Nekrutenko A, Hardison RC. 2004. Comparative genomics[J]. *Annu Rev Genomics Hum Genet*, **5**: 15-56.

Müller S, Wienberg J. 2001. "Bar-coding" primate chromosomes: molecular cytogenetic screening for the ancestral hominoid karyotype[J]. *Hum Genet*, **109**(1): 85-94.

Miyata T, Yasunaga T. 1980. Molecular evolution of mRNA: a method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application[J]. *J Mol Evol,* **16**(1): 23-36.

Navarro A, Barton NH. 2003a. Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation[J]. *Evolution Int J Org Evolution*, **57**(3): 447-459.

Navarro A, Barton NH. 2003b. Chromosomal speciation and molecular divergence--accelerated evolution in rearranged chromosomes[J]. *Science*, **300**(5617): 321-324.

Nielsen R, Bustamante C, Clark AG, Glanowski S, Sackton TB, Hubisz MJ, Fledel-Alon A, Tanenbaum DM, Civello D, White TJ, Sninsky JJ, Adams MD, Cargill M. 2005. A scan for positively selected genes in the genomes of humans and chimpanzees[J]. *PLoS Biol*, **3**(6): e170.

Normile D. 2001. Comparative genomics. Gene expression differs in human and chimp brains[J]. *Science*, **292**(5514): 44-45.

Pilpel Y, Sudarsanam P, Church GM. 2001. Identifying regulatory networks by combinatorial analysis of promoter elements[J]. *Nat Genet*, **29**(2): 153-159.

Roberts E, Hampshire DJ, Pattison L, Springell K, Jafri H, Corry P, Mannon J, Rashid Y, Crow Y, Bond J, Woods C. 2002. Autosomal recessive primary microcephaly: an analysis of locus heterogeneity and phenotypic variation[J]. *J Med Genet*, **39**(10): 718-721.

Samonte RV, Eichler EE. 2002. Segmental duplications and the evolution of the primate genome[J]. *Nat Rev Genet*, **3**(1): 65-72.

Schmitz J, Ohme M, Zischler H. 2001. SINE insertions in cladistic analyses and the phylogenetic affiliations of Tarsius bancanus to other primates[J]. *Genetics*, **157**(2): 777-784.

She X, Cheng Z, Zollner S, Church DM, Eichler EE. 2008. Mouse segmental duplication and copy number variation[J]. *Nat Genet*, **40**(7): 909-914.

She X, Horvath JE, Jiang Z, Liu G, Furey TS, Christ L, Clark R, Graves T, Gulden CL, Alkan C, Bailey JA, Sahinalp C, Rocchi M, Haussler D, Wilson RK, Miller W, Schwartz S, Eichler EE. 2004. The structure and evolution of centromeric transition regions within the human genome[J]. *Nature*, **430**(7002): 857-864.

She XW, Liu G, Ventura M, Zhao SY, Misceo D, Roberto R, Cardone MF, Rocchi M, Green ED, Archidiacano N, Eichler EE. 2006. A preliminary comparative analysis of primate segmental duplications shows elevated substitution rates and a great-ape expansion of intrachromosomal duplications[J]. *Genome Res*, **16**(5): 576-583.

Springer MS, Murphy WJ, Eizirik E, O'Brien SJ. 2003. Placental mammal diversification and the Cretaceous-Tertiary boundary[J]. *Proc Natl Acad Sci USA*, **100**(3): 1056-1061.

Stankiewicz P, Shaw CJ, Withers M, Inoue K, Lupski JR. 2004. Serial segmental duplications during primate evolution result in complex human genome architecture[J]. *Genome Res*, **14**(11): 2209-2220.

Stark A, Brennecke J, Bushati N, Russell RB, Cohen SM. 2005. Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution[J]. *Cell*, **123**(6): 1133-1146.

Szamalek JM, Goidts V, Searle JB, Cooper DN, Hameister H, Kehrer-Sawatzki H. 2006. The chimpanzee-specific pericentric inversions that distinguish humans and chimpanzees have identical breakpoints in *Pan troglodytes* and *Pan paniscus*[J]. *Genomics*, **87**(1): 39-45.

Taylor J, Tyekucheva S, Zody M, Chiaromonte F, Makova KD. 2006. Strong and weak male mutation bias at different sites in the primate genomes: insights from the human-chimpanzee comparison[J]. *Mol Biol Evol*, **23**(3): 565-573.

Trask BJ, Friedman C, Martin-Gallardo A, Rowen L, Akinbami C, Blankenship J, Collins C, Giorgi D, Iadonato S, Johnson F, Kuo WL, Massa H, Morrish T, Naylor S, Nguyen OT, Rouquier S, Smith T, Wong DJ, Youngblom J, van den Engh G. 1998. Members of the olfactory receptor gene family are contained in large blocks of DNA duplicated polymorphically near the ends of human chromosomes[J]. *Hum Mol Genet*, **7**(1): 13-26.

Trinklein ND, Aldred SJ, Saldanha AJ, Myers RM. 2003. Identification and functional analysis of human transcriptional promoters[J]. *Genome Res*, **13**(2): 308-312.

Vallender EJ, Lahn BT. 2004a. Effects of chromosomal rearrangements on human-chimpanzee molecular evolution[J]. *Genomics*, **84**(4): 757-761.

Vallender EJ, Lahn BT. 2004b. Positive selection on the human genome[J]. *Hum Mol Genet*, **13**(S2): R245-R254.

Wang XX, Grus WE, Zhang JZ. 2006. Gene losses during human origins[J]. *PLoS Biol*, **4**(3): e52.

Wang YQ, Su B. 2004. Molecular evolution of microcephalin, a gene determining human brain size[J]. *Hum Mol Genet*, **13**(11): 1131-1137.

Watanabe H, Fujiyama A, Hattori M, Taylor TD, Toyoda A, Kuroki Y, Noguchi H, BenKahla A, Lehrach H, Sudbrak R, Kube M, Taenzer S, Galgoczy P, Platzer M, Scharfe M, Nordsiek G, Blöcker H, Hellmann I, Khaitovich P, Pääbo S, Reinhardt R, Zheng HJ, Zhang XL, Zhu GF, Wang BF, Fu G, Ren SX, Zhao GP, Chen Z, Lee YS, Cheong JE, Choi SH, Wu KM, Liu TT, Hsiao KJ, Tsai SF, Kim CG, OOta S, Kitano T, Kohara Y, Saitou N, Park HS, Wang SY, Yaspo ML, Sakaki Y. 2004. DNA sequence and comparative analysis of chimpanzee chromosome 22[J]. *Nature*, **429**(6990): 382-388.

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS, Cook LL, Copley RR, Coulson A, Couronne O, Cuff J, Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dodge S, Dubchak I, Dunn DM, Eddy SR, Elnitski L, Emes RD, Eswara P, Eyras E, Felsenfeld A, Fewell GA, Flicek P, Foley K, Frankel WN, Fulton LA, Fulton RS, Furey TS, Gage D, Gibbs RA, Glusman G, Gnerre S,

Goldman N, Goodstadt L, Grafham D, Graves TA, Green ED, Gregory S, Guigó R, Guyer M, Hardison RC, Haussler D, Hayashizaki Y, Hillier LW, Hinrichs A, Hlavina W, Holzer T, Hsu F, Hua A, Hubbard T, Hunt A, Jackson I, Jaffe DB, Johnson LS, Jones M, Jones TA, Joy A, Kamal M, Karlsson EK, Karolchik D, Kasprzyk A, Kawai J, Keibler E, Kells C, Kent WJ, Kirby A, Kolbe DL, Korf I, Kucherlapati RS, Kulbokas EJ, Kulp D, Landers T, Leger JP, Leonard S, Letunic I, Levine R, Li J, Li M, Lloyd C, Lucas S, Ma B, Maglott DR, Mardis ER, Matthews L, Mauceli E, Mayer JH, McCarthy M, McCombie WR, McLaren S, McLay K, McPherson JD, Meldrim J, Meredith B, Mesirov JP, Miller W, Miner TL, Mongin E, Montgomery KT, Morgan M, Mott R, Mullikin JC, Muzny DM, Nash WE, Nelson JO, Nhan MN, Nicol R, Ning Z, Nusbaum C, O'Connor MJ, Okazaki Y, Oliver K, Overton-Larty E, Pachter L, Parra G, Pepin KH, Peterson J, Pevzner P, Plumb R, Pohl CS, Poliakov A, Ponce TC, Ponting CP, Potter S, Quail M, Reymond A, Roe BA, Roskin KM, Rubin EM, Rust AG, Santos R, Sapojnikov V, Schultz B, Schultz J, Schwartz MS, Schwartz S, Scott C, Seaman S, Searle S, Sharpe T, Sheridan A, Shownkeen R, Sims S, Singer JB, Slater G, Smit A, Smith DR, Spencer B, Stabenau A, Stange-Thomann N, Sugnet C, Suyama M, Tesler G, Thompson J, Torrents D, Trevaskis E, Tromp J, Ucla C, Ureta-Vidal A, Vinson JP, Von Niederhausern AC, Wade CM, Wall M, Weber RJ, Weiss RB, Wendl MC, West AP, Wetterstrand K, Wheeler R, Whelan S, Wierzbowski J, Willey D, Williams S, Wilson RK, Winter E, Worley KC, Wyman D, Yang S, Yang SP, Zdobnov EM, Zody MC, Lander ES. 2002. Initial sequencing and comparative analysis of the mouse genome[J]. *Nature*, **420**(6915): 520-562.

Webster MT, Smith NGC, Ellegren H. 2002. Microsatellite evolution inferred from human-chimpanzee genomic sequence alignments[J]. *Proc Natl Acad Sci USA*, **99**(13): 8748-8753.

Wetterbom A, Sevov M, Cavelier L, Bergström TF. 2006. Comparative genomic analysis of human and chimpanzee indicates a key role for indels in primate evolution[J]. *J Mol Evol*, **63**(5): 682-690.

Won YJ, Hey J. 2005. Divergence population genetics of chimpanzees[J]. *Mol Biol Evol*, 22: 297-307.

Yoder AD, Yang Z. 2000. Estimation of primate speciation dates using local molecular clocks[J]. *Mol Biol Evol* 17: 1081-90.

Yu N, Jensen-Seaman MI, Chemnick L, Kidd JR, Deinard AS, Ryder O, Kidd KK, Li WH. 2003. Low nucleotide diversity in chimpanzees and bonobos[J]. *Genetics*, **164**(4): 1511-1518.

Yu XJ, Zheng HK, Wang J, Wang W, Su B. 2006. Detecting lineage-specific adaptive evolution of brain-expressed genes in human using rhesus macaque as outgroup[J]. *Genomics*, **88**(6): 745-751.

Yunis JJ, Prakash O. 1982. The origin of man: a chromosomal pictorial legacy[J]. *Science*, **215**(4539): 1525-1530.

Zhang JZ. 2003. Evolution of the human ASPM gene, a major determinant of brain size[J]. *Genetics*, **165**(4): 2063-2070.

Zhang JZ, Nielsen R, Yang ZH. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level[J]. *Mol Biol Evol*, **22**(12): 2472-2479.

Zhang R, Peng Y, Wang W, Su B. 2007. Rapid evolution of an X-linked microRNA cluster in primates[J]. *Genome Res*, **17**(5): 612-671.